基于生成对抗网络的遮挡表情识别

王素琴，高宇豆，张加其

(华北电力大学 控制与计算机工程学院，北京 102206)

摘　要：针对实际应用中局部遮挡会影响人脸表情识别，提出一种基于生成对抗网络（GAN）的表情识别算法，先对遮挡人脸图像填补修复，再进行表情识别。其中 GAN 的生成器由卷积自动编码机构成，与鉴别器的对抗学习使得生成的人脸图像更加逼真。由卷积神经网络构成的鉴别器具有良好的特征提取能力，添加多分类层构成了表情分类器，避免了重新计算图像特征。为了解决训练样本不足的问题，将 celebA 人脸数据集用于训练人脸填补修复，同时表情分类器的特征提取部分得到了预训练。在 CK+数据集上的实验证明，填补后的人脸图像真实连贯，并取得了较高的表情识别率，尤其提高了人脸大面积遮挡的识别率。

关键词：人脸表情识别；局部遮挡；人脸修复；生成对抗网络；卷积神经网络

中图分类号：TP391.4　　doi: 10.3969/j.issn.1001-3695.2018.06.0360

Occluded facial expression recognition based on generative adversarial networks

Wang Suqin, Gao Yudou, Zhang Jiaqi

(*School of Control & Computer Engineering*, *North China Electric Power University*, *Beijing* 102206, *China*)

Abstract: Aiming at the fact that partial occlusion affects facial expression recognition in practical applications, an expression recognition method based on generative adversarial networks (GAN) is proposed. Firstly, the occlusion face images are filled and repaired, and then the expression recognition is performed. The generator of GAN is composed of a convolutional Auto-encoder, the face images generated by adversarial learning between generator and discriminator are more vivid. The discriminator is composed of the convolutional neural network and it has good feature extraction ability, and a multi-classification layer is added to construct the expression classifier, which avoids feature re-calculation. In order to solve the problem of insufficient training samples, the celebA face dataset is used to train face filling and repairing, and the feature extraction part of the expression classifier is pre-trained. Experiments on the CK+ dataset show that the face images after filling are real and coherent, and a higher expression recognition rate is achieved, especially the recognition rate of large-area occlusion of the face is improved.

Key words: facial expression recognition; partial occlusion; face completion; generative adversarial network; convolutional neural network

0　引言

人类情感多借助于面部表情进行传递[1]，利用图像处理技术对人脸面部表情进行识别具有重要意义。随着计算机技术的进步以及 GPU 等电脑硬件的发展，表情识别得到了长足的发展。文献[2~4]使用卷积神经网络 CNN 在标准的表情数据集上取得了很好的识别准确率。除了常见的 7 种基本表情，微表情持续时间很短，却饱含了人们隐藏的真实情感，对其进行识别是表情识别领域的一个研究热点，将微表情识别用于测谎，对公共安全、侦察破案等具有十分重要的作用[5]。但在实际应用中，拍摄到的人脸图像往往存在局部遮挡，常见的遮挡物有手、眼镜、口罩等，这些遮挡会干扰表情特征的提取，影响表情判别的准确性。

目前对局部遮挡的人脸进行表情识别一般采用非深度学习的方法，其主要思路总体上分为舍弃法和填补法两种。舍弃法是通过稀疏表示等方法将遮挡部分的信息简化或者丢弃，主要根据未遮挡部位进行表情识别。文献[6,7]基于稀疏表示提出了遮挡人脸表情识别方法，利用未遮挡部分构建字典，通过稀疏求解得到稀疏表示系数，最终实现待测人脸图像的表情判别。由于人的嘴巴、眼睛、鼻子等含有大量的表情信息，当这些部位被遮挡时，直接舍弃明显不合理。填补法则是先对遮挡部分进行填补，尽可能还原人脸未遮挡的状态，再进行表情识别。

收稿日期：2018-06-06；**修回日期：**2018-07-18

作者简介：王素琴（1970-），女，辽宁人，副教授，硕士，主要研究方向为图像处理，数据挖掘（wsq@ncepu.edu.cn）；高宇豆（1995-），女，硕士研究生，主要研究方向为计算机视觉、图像理解；张加其（1994-），男，硕士研究生，主要研究方向为计算机视觉、机器学习.

薛雨丽等人[8]采用鲁棒主成分分析法和显著性检测法重构被遮挡像素，然后由权值更新的 AdaBoost 分类器对去除遮挡的人脸图像进行表情识别。由于其使用的信息较为局限，难以保证填补效果，同时人工特征的设计繁琐，难以应对复杂的场景变化，鲁棒性较差。文献[9,10]利用卷积神经网络直接进行有遮挡的人脸表情识别，提高了识别的鲁棒性，但当遮挡区域较大时，准确率迅速下降。

因此，对于有遮挡特别是大面积遮挡的人脸图像，考虑先将其遮挡区域进行填补，使整幅图像看上去真实连贯，有利于表情判别。本文基于生成对抗网络构建一个遮挡表情识别模型，充分利用图像中未遮挡区域的像素信息，输入到由卷积自动编码机构建的生成器中进行人脸遮挡图像的填补修复，鉴别器在大量对抗训练中可以学习到人脸图像的特征表示，基于该特征表示添加多分类层构成分类器可以对填补后的图像进行表情识别。

1 相关工作

生成对抗网络（generative adversarial networks,GAN）[11]是当前人工智能领域的一大研究热点，是一种生成式模型，由生成器（generator，G）和鉴别器（discriminator，D）共同组成。生成器通过学习真实样本数据的分布规律来生成尽可能真实的伪数据，而鉴别器本质上为一个二分类器，它需要甄别出输入数据是真实样本还是生成器的输出。受二人零和博弈思想的启发，提出让生成器和鉴别器之间相互对抗，在对抗中迭代优化，不断提升各自的生成能力和鉴别能力，使生成器能够估测到数据样本的分布，生成的样本效果逼真。

GAN 模型在图像生成、图像修补以及图像去噪等领域应用广泛，同时衍生出许多改进的模型。文献[12]将 GAN 和卷积神经网络（CNN）结合起来得到深度卷积生成对抗网络（DCGAN），生成器和鉴别器均为卷积神经网络，在各个数据集上进行训练都得到了不错的结果，其特征表达能力强，使用得到的特征表示进行图像分类，效果较好。Brandon Amos 等人[13]直接使用该模型在人脸数据集上进行了图像填补实验，从模型生成的人脸图像中找到与原始图像最相似的一幅，然后根据其遮挡区域的位置掩码来替代填补原始图像，该方法可以填补得到一张内容上完整的人脸图像，但存在图像整体连贯性较差，适用性受限等问题。文献[14]提出了一个基于深度生成模型的人脸修复算法，它使用了两个鉴别器，分别判断整个图像区域和缺失区域的真假，最后使用一个解析网络进一步完善缺失区域的生成图像，该算法能够生成比较逼真的人脸图像，但将其结果直接用于表情识别存在诸多不适应，因为本文的目的是为了得到正确的表情类别，其填补内容需要在不误导表情判别的基础上，合理地增加表情信息。

除了生成视觉逼真的图像，GAN 还可以用于分类。深度神经网络在图像识别、对象分类等领域取得的成功很大程度上取决于大量的手动标注的训练数据集，然而在许多应用中，这样带标签的数据量往往不能满足深度模型训练的需要，将无标签的样本数据加入训练可进行半监督分类。文献[15,16]提出了将 GAN 的鉴别器由原来的二分类改为一个多分类器，此时生成器输出样本可用于训练分类器，作为 N 分类问题的第 N+1 类。该模型不仅利用有标签的训练样本，而且从无标签的生成样本中学习特征。本文对人脸进行填补修复以后还要完成一个表情多分类问题，因此基于这种半监督分类的思想，考虑将 GAN 的鉴别器和表情分类器结合起来，若使用深度卷积神经网络来构造 GAN 模型的鉴别器，其本身就具有很强的特征提取能力，可以用于多分类任务，同时生成器的输出样本也发挥了扩充样本数量的作用，提高分类模型的泛化能力。

2 本文模型

本文基于 GAN 构建了一个鲁棒的遮挡表情识别模型，对于一幅遮挡的人脸图像，首先要合成其遮挡区域的内容得到填补图像，进而正确地判断表情类别。模型的总体架构如图 1 所示，图中以黑色矩形框模拟遮挡物。

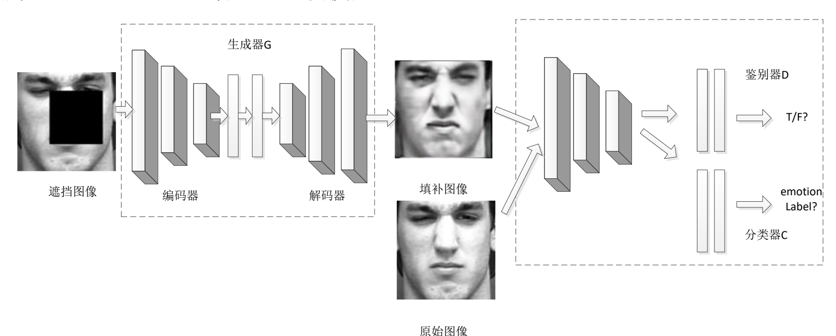


图 1 模型的总体架构

模型主要由生成器 G、鉴别器 D 和分类器 C 三个部分构成，功能上则可分为人脸填补和表情识别两个模块。人脸填补模块由生成器和鉴别器组成，将遮挡人脸图像输入到生成器中得到填补后的人脸图像，为了使填补效果更真实合理，通过鉴别器来判断图像为真实无遮挡人脸（真）还是生成器填补图像（假）。表情识别模块在鉴别器的基础上，使用了其部分卷积层和池化层，将其作为特征提取器，额外加入两层全连接层和 Softmax 层构成了表情分类器。

2.1 生成器 G

不同于原始 GAN，本文生成器 G 的输入不再是随机噪声，

而是遮挡人脸图像。其本质上是一个卷积自动编码机，由编码器和解码器组成，首先编码器通过多次卷积池化操作将模型输入映射到一个隐藏层表示，其中包含了原遮挡图像中已知区域和待填补区域之间的隐含关系，解码器利用这些隐含信息生成填补内容。生成器的网络结构如图 2 所示，其中编码器和解码器网络结构对称，互为逆操作。编码器参考 VGG[17]的模型结构，由卷积层 Conv()和池化层 Pooling()组成，其中 Conv(64，3×3，1×1)表示该卷积层由 64 个大小为 3×3，步长为 1 的卷积核组成，在每个卷积层后都紧接一个 LeakyReLU 激活层和一个 Batch normalization 归一化层，在池化层中进行最大池化操作，pooling(2×2)表示窗口大小为 2×2 的池化层 ，负责下采样。相应地，解码器通过卷积层 Conv()和上采样层 Upsampling()逐步恢复人脸图像，在编码器和解码器之间采用了两层神经元个数为 1024 的全连接层 Dense(1024)作为中间层。

2.2 鉴别器 D

鉴别器 D 对输入图像进行真假二分类。鉴别器的输入维度和生成器的输出维度一致，都为 128×128。在 DCGAN 中鉴别器为卷积核大小为 5×5 的全卷积网络，由于卷积核较大、网络层次不深导致其特征提取能力有限，若仅用于一般意义上的二分类是足够的，但为了将其特征提取部分和后续表情分类器的特征提取部分整合起来，需要一个网络更深、结构更好的深度卷积神经网络。其中 VGG16 的卷积核大小为 3×3，一共有 16 层，包括 13 个卷积层和 3 个全连接层，多个卷积层与非线性的激活层交替的结构使得其特征提取能力较强，同时模型深度、性能等方面也较适合，因此本文基于 VGG16 构建了鉴别器。

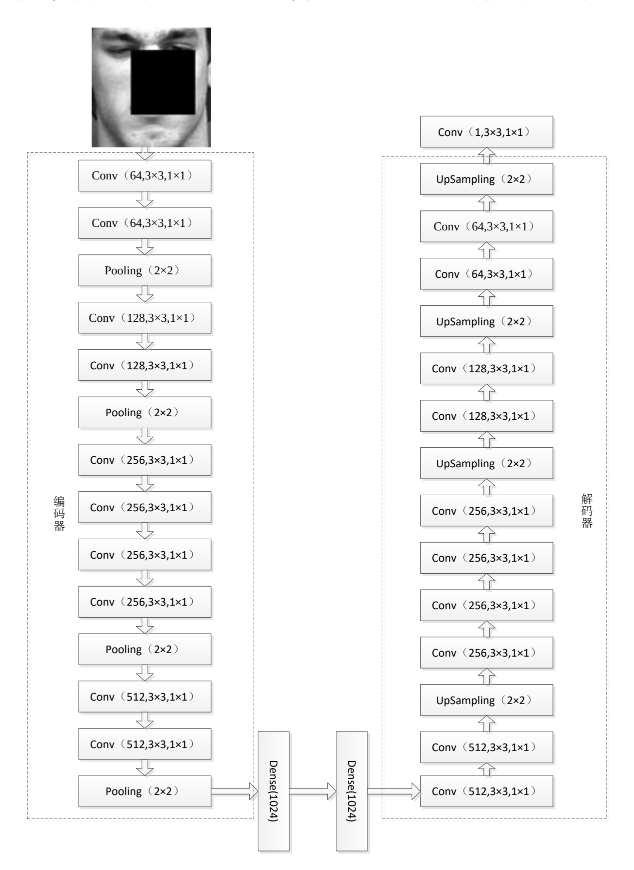


图 2　生成器网络结构

2.3 分类器 C

计算机为了完成表情识别任务，需要大量的带表情标签的训练数据，然而现有的表情数据集的数据有限，特别是将深度卷积神经网络引入后，更是需要大量的数据支撑，如果数据太少模型在训练中会很容易达到过拟合不能发挥模型本身的能力。CK+[18]数据集 (The Extended Cohn-Kanade AU-Coded Database)是标准的人脸表情数据集，其中仅含有 327 个带标签的峰值表情图像，即使将部分非峰值表情图像也加入训练仍然不够。不过在人脸填补模块，模型的训练并不需要表情标签，因此可以利用现有的大量的人脸图像数据进行训练，鉴别器在训练过程不断加强人脸特征提取能力，利用鉴别器得到的特征表示可进行表情分类。因此分类器共享了鉴别器的部分卷积层和池化层来提取特征，总体上由特征提取层、两层全连接层和 Softmax 分类层构成，本质上是一个卷积神经网络。

2.4 损失函数

首先为生成器定义了重构损失函数 l_r ，它是通过计算生成网络输出结果 G(z)和原始未遮挡图像 y 对应位置像素点 i 的像素差值来得到的，本文通过 L2 范数来衡量，值越小说明相似度越高。 l_r 定义为

$$l_r=\sum_{i=0}^{m}(y^i-G(z)^i)^2 \tag{1}$$

为了让生成的图像尽可能真实，本文引入了对抗损失 l_a，生成器 G 希望输出图像 G(z)能够迷惑鉴别器 D，而鉴别器则希望能准确判断输入图像的真假。和原始 GAN 对抗损失的定义一样， l_a 定义为如式（2）所示，Pdata 为标准人脸图像的数据分布， p_z 是遮挡的人脸图像的数据分布。

$$\begin{aligned} l_a = \min_G \max_D\ & \varepsilon_{x\sim p_{data}(x)}[\log D(x)] + \\ & \varepsilon_{x\sim p_z(z)}[\log(1-D(G(z)))] \end{aligned} \tag{2}$$

在分类器中，采用交叉熵损失 l_c 来训练表情多分类，交叉熵描述了预测的概率分布 q(x)和真实概率分布 p(x)之间的距离。 l_c 定义为

$$l_c=-\sum_x p(x)\log q(x) \tag{3}$$

模型的总体损失函数 L 由 l_r 、 l_a 和 l_c 三部分构成，定义为

$$L=l_r+\lambda_1 l_a+\lambda_2 l_c \tag{4}$$

其中: λ_1 和 λ_2 是权重因子，用于平衡二分类和多分类损失的比重。

3 遮挡表情识别

本文实验在 16GB 内存的 NVIDIA GeForce GTX 1080Ti 的 GPU 工作站上进行，基于标准人脸表情数据集 CK+来衡量模型对遮挡人脸的表情识别能力。由于 CK+数据量较小，需先利用 CelebA[19]人脸数据集对人脸填补模块生成器和鉴别器进行预训练，再加入分类器在 CK+数据集上进一步微调。

3.1 数据集处理

CK+数据集包括 123 个人共 595 例的表情序列，包含从中性到表情峰值的所有图像帧，其中仅有 327 例图像序列有明确的标签，共分为 7 种表情（悲伤、快乐、愤怒、厌恶、惊讶、蔑视和恐惧）。为了扩大样本量，同时考虑实用性，加入非峰值表情以增强模型的泛化能力。将每例图像序列的开始图像作为中性表情样本，将图像序列中接近表情峰值的若干张图像作为该表情图像样本。共得到 3559 张表情样本，其中 876 张中性表情，260 张悲伤表情，627 张快乐表情，482 张愤怒表情，417 张厌恶表情，640 张惊讶表情，257 张恐惧表情。

为了进一步扩大样本量，并解决各类表情样本数量不均衡的问题，对数据集中的数据进行镜像翻转、旋转等操作，扩展后共有 10537 张表情图像，各类表情图像的数量在 1200 张左右。随机选择其中的 1/5 作为测试集，其余 4/5 为训练集，同时满足测试集和训练集的人物信息不交叉。

CelebA 人脸数据集来源于香港中文大学，包含 10177 个名人的 202599 张人脸图像。选取裁剪出人脸部分的图像作为本文基础训练数据集，随机选择其中的 4/5 作为训练集，剩余的 1/5 数据用于测试模型生成器的填补能力以及鉴别器的真假判别能力。

3.2 图像预处理及遮挡模拟

为了尽量避免光照和姿态变换等因素对表情识别的干扰，需要做图像预处理。首先对人脸图像进行归一化处理，采用加权平均法对彩色人脸图像进行灰度化处理，随后进行直方图均衡化处理，在不损失图像信息的基础上，降低图像数据大小，减少光照因素的干扰。根据检测出的人脸关键点（通常在眼角、鼻子、嘴巴、脸轮廓的位置）进行人脸对齐校准，可消除头部姿势不同对后续操作带来的误差，同时人脸各个器官部位在相对固定的位置将有助于特征的提取和分析。本文使用了 OpenFace[20]的人脸对齐算法，采用集成回归树（ERT）[21]估计面部特征点来对齐人脸图像，最终校准后统一将图像的大小调整为 128×128。

现实生活中一般存在两种遮挡，一种是由于手或者头部运动产生的临时遮挡，另一种则是由墨镜、口罩、围巾等引起的系统遮挡[22]。由于目前没有一个通用的、成熟的遮挡人脸表情数据集，因此本文通过在人脸图像的不同位置添加黑色矩形框来模拟眼部、嘴部和随机遮挡，分别对应眼镜、口罩引起的系统遮挡和临时遮挡。经过图像预处理和遮挡模拟后的人脸图像如图 3 所示。经过人脸对齐操作，图像中的面部器官基本在同一位置，对图像中固定位置进行遮挡即可简单模拟系统遮挡，其中眼部遮挡面积大约为整个图像的 15%，嘴巴遮挡面积大约为整个图像的 25%。另外随机遮挡 10%、20%、30%、40%、50%、60%用于模拟位置不固定的临时遮挡，本文认为遮挡面积超过 60%时，表情识别意义不大，故不作分析。

3.3 未填补前表情分类器的对比实验

为了进一步说明对遮挡人脸图像进行填补的意义以及和本文模型进行对比，本文在填补修复遮挡人脸图像前，直接使用了卷积神经网络对遮挡人脸图像进行了实验，同样采用了 VGG16。但可供训练的数据量较少、网络较深导致模型从头训练难以达到很好的效果。VGGface 的权重是在 LFW（labeled faces in the wild）数据集[23]上训练得到的，可用于与人脸图像相关的训练。基于迁移学习的思想，本文使用了 VGGface 的权重，在 CK+数据集上进行微调，能够快速收敛且不需要很大的数据量。实验中采用了两种方法来训练卷积神经网络，方法 1 仅使用了标准的人脸图像进行训练，方法 2 则将遮挡的人脸图像也加入了模型的训练。


无遮挡


眼部遮挡


嘴部遮挡

随机遮挡 10%


随机遮挡 20%


随机遮挡 30%

随机遮挡 40%

随机遮挡 50%

随机遮挡 60%

图 3 图像预处理和遮挡模拟

3.4 本文模型训练

为了有效地训练模型，将整个模型的训练过程分为两大部分，分别是对人脸填补网络和表情分类网络的训练。模型首先在 celebA 数据集上训练人脸填补网络，先使用重构损失训练生成器使其得到一个模糊的填补内容，接着加入鉴别器对抗损失（即二分类损失）来联合训练，模型使用 Adam 优化器，初始学习率为 0.0002，剩余的 Adam 超参数设置为默认值，50 次迭代后学习率下降一个量级。当 GAN 模型接近纳什均衡，二分类准确率大约 0.5 时，鉴别器几乎不能判断输入样本的真假，此时不再使用 CelebA 数据集，而是在 CK+数据集上进行微调，同时加入了表情分类器的训练，起初通过设置样本权重，让模型使用填补图像和无遮挡图像各一半训练二分类器，而仅使用无遮挡图像训练多分类器，经过 30 次迭代后，将填补图像也加入到多分类器的训练。

4 实验结果及分析

4.1 未填补前表情分类器的实验结果

经 4.3 中两种方式训练 VGGface 后，对不同类型的遮挡人脸图像都进行了测试，结果如表 1 所示。

表 1 基于 VGGface 的表情识别实验结果

遮挡类型	遮挡面积	方法 1	方法 2
无遮挡	0	97.23	97.57
随机遮挡 10%	40×40	80.46	94.98
随机遮挡 20%	60×60	59.93	92.75
随机遮挡 30%	70×70	49.71	89.69
随机遮挡 40%	81×81	42.56	77.51
随机遮挡 50%	90×90	25.69	67.57
随机遮挡 60%	99×99	15.53	51.40
眼部遮挡（15%）	25×95	89.90	92.69
嘴部遮挡（25%）	40×100	48.24	87.34

可以看出，对于无遮挡的人脸图像，卷积神经网络能够很好地提取特征并进行分类，但随着人脸遮挡面积的增加，表情识别准确率锐减。将遮挡人脸加入训练，实验发现将遮挡面积较小的人脸图像加入训练，相当于进行了数据增强，模型整体的识别准确率上升，但当将大面积遮挡的人脸图像也加入训练时，遮挡面积较小的人脸图像的表情识别准确率会有所下降，因此，本文仅使用了遮挡面积小于等于 30%的人脸图像用于训练，最终遮挡人脸图像的测试准确率有所上升，但与无遮挡的表情识别准确率有较大差距，仍需进一步提高，尤其是遮挡面积大于 30%的时候。

4.2 本文方法实验结果分析

本文模型的填补效果如图 3 所示。遮挡类型从左到右分别是随机遮挡 10%，随机遮挡 20%，随机遮挡 40%，眼部遮挡以及嘴部遮挡，第 1 列和第 4 列是中性表情，第 2、3、5 列依次是蔑视、惊讶和高兴的表情。其中第一行是遮挡前的原始图像，第二行是遮挡后的人脸图像，第三行是人脸填补修复后的结果。



图 3 遮挡人脸填补结果

由图 3 可以看出，经过人脸填补修复，整个人脸图像看上去真实连贯，未遮挡区域可能发生细微变化，但一般不会影响人脸图像的表情类别。

将填补后的人脸图像输入到表情分类网络中，对表情进行判别，实验结果如图 4 所示。为了对比填补前后人脸表情识别效果，在填补前采用了 4.1 节中的方法 2 直接对遮挡人脸图像进行了表情识别。可以看出，人脸填补模型整体上提升了遮挡表情识别的准确率，当遮挡面积小于 60%时，模型整体识别率还在 80%以上，较方法 2 直接使用卷积神经网络提升了 28.96%，而在遮挡面积较小的情况下，模型的识别准确率提升幅度较小。考虑模型的处理时间、系统资源等方面因素，在遮挡面积小于面部面积 10%时，可不进行填补，直接使用卷积神经网络进行识别。

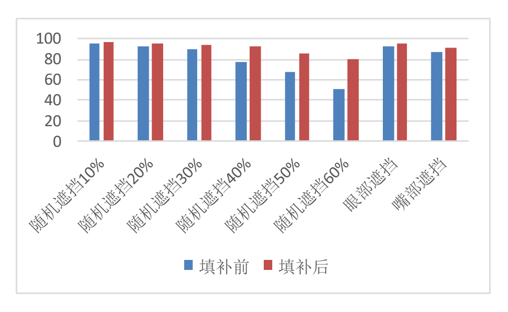


图 4 遮挡人脸表情识别的对比实验结果

4.3 对比其他方法

本文选取了文献[24]的 PCA+SVM 方法、文献[6]的稀疏表示方法（sparse representation based classification, SRC）以及 CNN、DCGAN[13]+CNN 方法来进行对比，其中 CNN 的方法即 4.1 中的方法 2，DCGAN+CNN 方法中 DCGAN 用于填补遮挡人脸图像，CNN 同为微调的 VGGface 模型，用于人脸表情分类。所有方法都在 CK+数据集上进行了实验，结果如表 2 所示。可以看出，无论人脸图像是否存在遮挡，本文方法的表情识别准确率都较高，深度学习方法中 DCGAN+CNN 方法虽然也对人脸图像进行了填补，但得到的图像连贯性较差，影响了表情识别的准确率，尤其是遮挡面积小于 40%，该方法的表情识别准确率甚至小于仅用 CNN 的准确率。而本文的方法避免了图像不连贯对表情识别的影响，在人脸遮挡面积 60%时，仍能达到 80%以上的识别准确率。

表 2 与其他方法对比实验结果

方法	无遮挡	随机遮挡 10%	随机遮挡 20%	随机遮挡 30%	随机遮挡 40%	随机遮挡 50%	随机遮挡 60%
PCA+SVM	91.23	89.32	85.93	84.13	79.23	68.87	54.97
SRC	92.79	90.23	88.69	85.97	80.00	70.23	55.78
CNN	97.57	94.98	92.75	89.69	77.51	67.57	51.40
DCGAN+CNN	97.57	89.30	86.33	83.93	78.94	70.91	61.92
本文模型	97.57	96.45	95.11	93.34	92.02	85.89	80.34

5 结束语

实际应用中局部遮挡问题导致表情识别的准确率不够理想，与在标准人脸表情数据集上取得的高识别率差距明显。本文提出了一种基于 GAN 的遮挡表情识别模型，对人脸图像中遮挡缺失的部分先进行填补修复，再基于卷积神经网络进行表情识别。将遮挡人脸图像作为输入条件，由卷积自动编码机构成的

生成器输出一张完整的人脸图像，在与鉴别器的相互对抗中，生成的人脸图像更加逼真自然。将鉴别器和分类器的特征提取部分整合在一起，利用常规的人脸图像数据集 celebA 对其权重参数进行预训练，最后在 CK+数据集上进行微调得到了很好的分类结果。经实验，本文方法填补的人脸图像直观上真实连贯，填补后的表情识别率较其他方法更高。模型尤其提高了大面积遮挡人脸图像的表情识别准确率。

参考文献：

[1] Mehrabian Albert, Russell James A. An approach to environmental psychology [M]. Cambridge: MIT Press, 1980: 222-253.

[2] 孙晓，潘汀，任福继．基于 ROI-KNN 卷积神经网络的面部表情识别 [J]. 自动化学报, 2016, 42 (6): 883-891. (Sun Xiao, Pan Ting, Ren Fuji. Facial expression recognition using ROI-KNN deep convolutional neural networks [J]. Acta Automatica Sinica, 2016, 42 (6): 883-891.)

[3] 李江，冉君军，张克非．一种基于降噪自编码器的人脸表情识别方法 [J]. 计算机应用研究, 2016, 33 (12): 3843-3846. (Li Jiang, Ran Junjun, Zhang Kefei. Method of facial expression recognition based on denoising autoencoders [J]. Application Research of Computers, 2016, 33 (12): 3843-3846.)

[4] Lopes A T, De Aguiar E, Oliveira-Santos T. A facial expression recognition system using convolutional networks [C]// Proc of the 28th SIBGRAPI Conference on Graphics, Patterns and Images. 2015: 273-280.

[5] 贲晛烨，杨明强，张鹏，等．微表情自动识别综述 [J]. 计算机辅助设计与图形学学报, 2014, 26 (9): 1385-1395. (Ben Xian ye, Yang Mingqiang, Zhang Peng, *et al.* Survey on automatic micro expression recognition methods [J]. Journal of Computer-Aided Design & Computer Graphics, 2014, 26 (9): 1385-1395.)

[6] 朱明旱，李树涛，叶华．基于稀疏表示的遮挡人脸表情识别方法 [J]. 模式识别与人工智能, 2014, 27 (8): 708-712. (Zhu Minghan, Li Shutao, Ye Hua. An occluded facial expression recognition method based on sparse representation [J]. PR & AI, 2014, 27 (8): 708-712.)

[7] Cotter S F. Recognition of occluded facial expressions using a fusion of localized sparse representation classifiers [C]// Proc of Digital Signal Processing Workshop and IEEE Signal Processing Education Workshop. 2011: 437-442.

[8] 薛雨丽，毛峡，Catalin-Daniel C，等．遮挡条件下的鲁棒表情识别方法 [J]. 北京航空航天大学学报, 2010, 36 (4): 429-433. (Xue Yuli, Mao xia, Catalin-Daniel C, *et al.* Robust facial expression recognition under occlusion condition. Journal of Beijing University of Aeronautics and Astronautics, 2010, 36 (4): 429-433.)

[9] 王军，夏利民．基于深度学习的痛苦表情识别 [J]. 计算机工程与设计, 2016, 37 (6): 1617-1620. (Wang Jun, Xia Limin. Pain expression recognition based on deep learning [J]. Computer Engineering and Design, 2016, 37 (6): 1617-1620.)

[10] 王剑云，李小霞．一种基于深度学习的表情识别方法 [J]. 计算机与现代化, 2015 (1): 84-87. (Wang Jianyun, Li Xiaoxia. A facial expression recognition method based on deep learning [J]. Computer and Modernization, 2015 (1): 84-87.)

[11] Goodfellow l, Pouget-Abadie J, Mirza M, *et al.* Generative adversarial nets [C]// Proc of International Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2014: 2672-2680.

[12] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks [EB/OL]. (2016-01-07) . https://arxiv. org/abs/1511. 06434.

[13] Yeh R, Chen Chen, Lim T Y, *et al.* Semantic Image Inpainting with Perceptual and Contextual Losses [C]// Proc of IEEE Conference on Computer Vision and Pattern Recognition. 2017: 6882-6890.

[14] Li Yijun, Liu Sifei, Yang Jimei, *et al.* Generative face completion [C]// Proc of IEEE Conference on Computer Vision and Pattern Recognition. 2017: 5892-5900.

[15] Salimans T, Goodfellow I, Zaremba W, *et al.* Improved techniques for training GANs [C]// Proc of Neural Information Processing Systems Conference. 2016.

[16] Denton E, Gross S, Fergus R. Semi-supervised learning with context-conditional generative adversarial networks [EB/OL]. (2016-11-19) . https://arxiv. org/abs/1611. 06430.

[17] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition [EB/OL]. (2015-04-10) . https://arxiv. org/abs/1409. 1556.

[18] Lucey P, Cohn J F, Kanade T, *et al.* The extended Cohn-Kanade dataset (Ck+): a complete dataset for action unit and emotion-specified expression [C]// Proc of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington DC: IEEE Computer Society, 2010: 94-101.

[19] Yang Shuo, Luo Ping, Loy C C, *et al.* From facial parts responses to face detection: a deep learning approach [C]// Proc of IEEE International Conference on Computer Vision. 2015: 3676-3684.

[20] Baltrušaitis T, Robinson P, Morency L P. Openface: an open source facial behavior analysis toolkit [C]// Proc of IEEE Winter Conference on Applications of Computer Vision. Piscataway, NJ: IEEE Press, 2016: 2016: 1-10.

[21] Kazemi V, Sullivan J. One millisecond face alignment with an ensemble of regression trees [C]// Proc of IEEE Conference on Computer Vision and Pattern Recognition. Washington DC: IEEE Computer Society, 2014: 1867-1874.

[22] Towner H, Slater M. Reconstruction and Recognition of occluded facial expressions using PCA [M]// Affective Computing and Intelligent Interaction. Berlin: Springer, 2007: 36-47.

[23] Huang G B, Mattar M, Berg T, *et al.* Labeled faces in the wild: a database for studying face recognition in unconstrained environments [R]. University of Massachusetts Amherst Technical Report, 2008: 07-49.

[24] 罗元，吴彩明，张毅．基于 PCA 与 SVM 结合的面部表情识别的智能轮

椅控制 [J]. 计算机应用研究, 2012, 29 (8): 3166-3168. (Luo Yuan, Wu Caiming, Zhang Yi. Facial expression recognition based on principal component analysis and support vector machine applied in intelligent wheelchair [J]. Application Research of Computers, 2012, 29 (8): 3166-3168.)